DOCUMENT RESUME

ED 361 355 TM 020 413

AUTHOR Qian, Gaoyin

TITLE Methodological Issues and Further Research in the Study of Reading Comprehension with Different Levels of Knowledge.

PUB DATE Apr 93

NOTE 18p.; Paper presented at the Annual Meeting of the American Educational Research Association (Atlanta, GA, April 12-16, 1993).

PUB TYPE Information Analyses (070) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.

DESCRIPTORS Age Differences; Data Collection; Educational Research; *Effect Size; Elementary Secondary Education; *Experiments; *Knowledge Level; Meta Analysis; *Reading Comprehension; Research Methodology; Research Needs; *Research Problems; Statistical Analysis; *Validity

ABSTRACT

Some methodological issues in the study of levels of knowledge are reviewed, and needs for further research are explored, drawing on an analysis of 12 studies reported since the late 1970s. In the 12 studies, 16 quantitative experiments were conducted. These were assessed for internal and external validity. Analysis revealed some shortcomings in study design, some confounds in data collection, some threats to statistical procedures, and other issues related to external validity. For example, in 11 of 12 experiments, subjects were not randomly selected. None of th 16 experiments included a discussion of data assumptions for the statistical procedures researchers chose to use. In 8 of the 16 experiments, no information was given about the number of subjects assigned to each group. Several other issues were found that might have affected validity. In addition, none of the studies reported the effect size of significant results. Research in more natural settings is required, and further study of the compensatory effects of prior knowledge is required. Research into age differences and on the effects of prior instruction is needed. Two tables present study data. (SLD)

Methodological Issues and Further Research in the Study of Reading Comprehension with Different Levels of Knowledge

Gaoyin Qian

309 Aderhold Hall

Department of Reading Education

University of Georgia

Athens, Georgia 30602

Paper presented at the annual meeting of the American Educational Research Association, Atlanta, Georgia, April 12, 1993

The development of cognitive psychology and the burgeoning of interest in students' knowledge organization have combined to offer insights to the reading process. Many cognitive theorists and researchers have demonstrated that readers' comprehension is enhanced if their preexisting knowledge is activated or if they are provided with opportunities to build background knowledge (Anderson, Spiro, & Anderson, 1978; Rumelhart & Ortony, 1977).

Since the late 1970s, new trends have emerged in the study of the acquisition of knowledge. Some researchers have investigated the effects of prior knowledge upon learning and comprehension of the readers with different levels of knowledge about passage topics, with different abilities, or with different levels of expertise (Chi, Glaser, & Rees, 1982; Means & Voss, 1985; Recht & Leslie, 1988; Schneider & Korkel, 1989; Schneider , Korkel, & Weinert, 1989; Spilich, Vesonder, Chiesi, & Voss, 1979; Stahl, Jacobson, Davis, & Davis, 1989). The results of this research have shown that the extent of knowledge has the effect on the quantity as well as quality of students' understanding of the text.

Efforts have also been made to explore the different levels of domain-specific knowledge and their relation to the acquisition of domain-specific knowledge. Research literature on levels of knowledge has revealed the superiority of high knowledge (HK) individuals (experts) over low knowledge (LK) individuals (novices) in certain aspects. HK individuals tend to outperform LK individuals by recalling more text information, providing rule-governed protocols (deep structures), engaging

in metacognitive processing and being more accurate in their solutions to the problems.

There is, however, scarcity of critical review of the studies in which the significant results were found in the investigation of levels of domain knowledge. In particular, there is a need to examine the methodological rigor of the studies so that research consumers can employ the useful information to evaluate and judge the quality of research in this field.

Therefore, it is the aim of this paper to address some of the methodological issues in the study of levels of knowledge and discuss the further research to be needed based on what has been found in the literature.

Results of the Study

The following discussion is based on an analysis of 12 studies reported in journal articles since the late 1970s. In the 12 studies, 16 quantitative experiments were conducted. Critical criteria to assess the experimental studies developed by Lysynchuk, Pressley, d'Ailly, Smith, and Cake (1989) were used to evaluate the internal and external validity of the quantitative experiments. The analysis revealed a shortcoming in the design of the studies, some confounds in data collection, some threats to the statistical procedures, and issues related to the external validity.

Insert Table 1 about Here

In general design, the most striking shortcoming is that in 11 out of 12 experiments (about 92%) the subjects were not randomly selected. This obvious shortcoming certainly affected the generalizability of the results found in the studies. What is more important is that lack of randomization may cause systematic bias on certain variables which may have attributed to the significant treatment effects.

In addition to lack of randomization in general design, components of the data collection process may have affected the results and conclusions of the studies we examined. There were some possible confounds that may threaten the internal validity of the studies. For example, 10 out of 16 experiments (about 63%) provided the information that subjects in all the groups were exposed to the experimental materials in the same amount of time. Only 9 out of 16 experiments (56%) reported the amount of time that subjects in each group spent on the dependent variable task.

Shortcomings in data collection also include lack of manipulation checks and process measures. The former refers to checks to make sure whether the subjects perform the tasks as directed, which occurred in 4 out of 16 experiments (25%). The latter emphasizes relatively direct measure or processing in addition to outcome measures, which occurred in 3 out of 16 experiments (about 19%). In a fairly high rate of the experiments (8 out of 13, about 62%), the researchers did not include interrater reliability in scoring the subjects' recall protocols, the most common technique used in the study of levels of domain knowledge.

In addition to threats to internal validity imposed by design and data collection,

issues related to statistical analysis may pose additional threats. The appropriateness of the statistical tests was determined by the following principles: (a) selection of the test statistic should logically fit the purpose of the study and best answer the questions that are being explored; (b) data assumptions of the selected test statistics should be met; (c) if assumptions are violated, researchers should be aware of the issues involved and resort to alternatives to deal with the problem. In the 16 experiments we examined, none included a discussion about data assumptions for the statistical procedure that researchers chose to use. In 8 out of 16 experiments (50%), no information was provided about the number of subjects assigned to each group, whereas in 6 out of 13 experiments (about 46%), the data of standard deviations were not included in the report. Without the information of the standard deviations we can hardly determine the data assumptions that some test statistics are based on. For example, in one study, a number of pooled t-tests were used to assess the difference between good and poor learners on several measures. Equal variance of the groups is considered to be an important assumption for such a test. However, the researchers did not give any information about variances. In another study the analysis of covariance was used to process the data; however, there was no evidence on the part of the researchers indicating that they tested linear relation between the dependent measure and covariate, and homogeneity of the regression slopes. Also, in one study the researchers were not explicit about the kind of test statistic used to process the data.

External validity concerns about generalizability of the results obtained to other situations. Two broad types of external validity have been described: (a) population validity and (b) ecological validity (Huck, Cormier, & Bounds, 1974). Population validity involves generalization to other populations. Ecological validity involves generalization to environments similar to those of the experimental conditions. While threats to internal validity are functions of research design, threats to external validity are not. In part, external invalidity is a result of inadequate descriptions of subjects, independent, and dependent variables (Huck et al., 1974). Since the late 1980s, there has been an increased interest in evaluating ecological validity in educational research. The same principles are applicable in the study of levels of domain knowledge.

Insert Table 2 about here

Table 2 summarizes the number (and percentage) of all studies that met a particular external validity criterion. A most disturbing problem is that researchers did not give ample description of the subjects they used in the experiments (13 out of 16, 81%). Besides, most of the researchers (15 out 16, about 94%) were not concerned about levels of reading ability of the subjects and the difficulty level of the materials they used in the study (13 out of 16, about 81%). Besides, the biggest problems found in the review in terms of ecological validity are the laboratory treatment condition and highly contrived text used in the experiments.

In addition to the above issues, some other problems that arose in the studies need to be addressed. In one study, a disproportionate number of female subjects were grouped in the low knowledge group. It is suspected that sex difference may have been attributed to the difference found between high knowledge (HK) and low knowledge (LK) groups because most females might be less interested in baseball games.

Another issue is about the use of large sample size found in two studies. The researchers could have included the effect size of the test so that research consumers can evaluate the significant findings from practical perspectives. By the way, none of the studies reported the effect size of the significant results.

Conclusion

This critique of the internal validity of reading comprehension research with different levels of knowledge by no means indicates that we are looking for "perfect" studies conducted in various settings. Instead, we argue that researchers should consider issues involved in statistical analyses when assumptions are not met, when unequal cell sizes are used in factorial design, and when "... a particular solution is selected on rational grounds so that those selections can be rationally described and defined" (Levin, 1985, p. 227). Statistical problems can be avoided by including information about means, standard deviations, and number of subjects.

When assumptions underlying statistical procedures are not met, there are alternative analysis strategies. In the case of repeated designs, when the assumption

of sphericity is not met, a univariate contrast approach might be appropriate (Levin, 1985). When the assumption of homogeneity of regression slopes is not tenable, the Johnson-Neyman technique is a viable alternative (Stenvens, 1990).

The extent to which results may be generalized to other populations depends in part on the degree of description of the sample participating in the study and the conditions under which the study is conducted. Researchers in comprehension with different levels of knowledge could increase the external validity of their work by providing more complete descriptions of dependent measures (inducing reliabilities), reading levels of materials, and reading abilities of students. External validity could improved further by including delayed measures. Determining long term effects of strategy instruction provides indications of internalization and may be useful in informing future instruction.

Future Research Needed

First, research is needed to extend some findings in the study of levels of knowledge to more natural settings by using natural text in different knowledge domains. The findings in some studies were based on more controlled and artificial text (Spilich et al., 1979; Chiesi, Spilich, & Voss, 1979). The results may be different in a more natural setting with a more authentic text.

Second, the efficacy of proving prior knowledge to compensate for inefficiency of low knowledge individuals and low aptitude is inconclusive. Further research is needed to specify where and in what situation the compensation does not occur.

Third, it remains a question to be answered by further research whether young experts could learn a more complex schema. There is an controversy over whether young experts would perform as older experts even if they could learn a more complex schema.

Finally, there is scarcity of intervention studies that have investigated the effect of instruction of prior knowledge and learning strategies on students' learning and comprehension in knowledge domain.

References

Anderson, R. C., Spiro, R. J., & Anderson, M. C. (1978). Schemata as scaffolding for the representation of information in connected discourse. American Educational Research Journal, 15(3), 433-440.

Chi, M. T. H., Glaser, R., & Rees, E. (1982). Expertise in problem solving. In R. Sternberg (Ed.), Advances in the psychology of human intelligence (Vol. 1) (pp. 7-75). Hillsdale, NJ: Erlbaum.

Chiesi, H. L., Spilich, G. J., & Voss, J. F. (1979). Acquisition of domain-related information in relation to high and low domain knowledge. Journal of Verbal Learning and Verbal Behavior, 18, 257-273.

Levin, J. R. (1985). Some methodological and statistical "Bugs" in research on children's learning. In M. Pressley & C. J. Brainerd (Eds.), Cognitive learning and memory in children: Progress in cognitive development research (pp. 205-233). New York: Springer-verlag.

Lysynchuk, L. M., Pressley, M., d'Ailly, H., Smith, M., & Cake, H. (1989). A methodological analysis of experimental studies of comprehension strategy instruction. Reading Research Quarterly, 24, 458-470.

Means, M. L., & Voss, J. F. (1985). Star wars: A developmental study of expert and novice knowledge structures. Journal of Memory and Language, 24, 746-757.

Recht, D. R., & Leslie, L. (1988). Effect of prior knowledge on good and poor readers' memory of text. Journal of Educational Psychology, 80, 16-20.

Rumelhart, D. C., & Ortony, A. (1977). The representation of knowledge in memory. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), Schooling and the acquisition of knowledge. Hillsdale, N. J.: Lawrence Erlbaum Associates.

Schneider, W., & Korkel, J. (1989). The knowledge base and text recall: Evidence from a short-term longitudinal study. Contemporary Educational Psychology, 14, 382-393.

Schneider, W., Korkel, J., & Weinert, F. (1989). Domain-specific knowledge and memory performance: A comparison of high-and low-aptitude children. Journal of Educational Psychology, 81(3), 306-312.

Spilich, G. J., Vesonder, G. T., Chiesi, H. L., & Voss, J. F. (1979). Text processing of domain-related information for individuals with high and low domain knowledge. Journal of Verbal learning and Verbal Behavior, 18, 275-290.

Stevens, J. P. (1990). Intermediate statistics: A modern approach. Hillsdale, NJ: Lawrence Erlbaum.

Stahl, A. S., Jacobson, M. G., Davis, C. E., & Davis, R. L. (1989). Prior knowledge and difficult vocabulary in the comprehension of unfamiliar text. Reading Research Quarterly, 24(1), 27-43.

Studies Evaluated

Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. Cognitive Science, 13, 145-182.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. Cognitive Science, 5, 121-152.

Chiesi, H. L., Spilich, G. J., & Voss, J. F. (1979). Acquisition of domain-related information in relation to high and low domain knowledge. Journal of Verbal Learning and Verbal Behavior, 18, 257-273.

Goldman, S. R., & Duran, R. P. (1988). Answering questions from oceanography texts: Learner, task, and text characteristics. Discourse Processes, 11, 373-412.

Means, M. L., & Voss, J. F. (1985). Star wars: A developmental study of expert and novice knowledge structures. Journal of Memory and Language, 24, 746-757.

Recht, D. R., & Leslie, L. (1988). Effect of prior knowledge on good and poor readers' memory of text. Journal of Educational Psychology, 80, 16-20.

Reynolds, R. E., & Taylor, M. A., Steffensen, M. S., Shirey, L. L., & Anderson, R. C. (1982). Cultural schemata and reading comprehension. Reading Research Quarterly, 17, 353-366.

Schneider, W., & Korkel, J. (1989). The knowledge base and text recall: Evidence from a short-term longitudinal study. Contemporary Educational Psychology, 14, 382-393.

Schneider, W., Korkel, J., & Weinert, F. (1989). Domain-specific knowledge and memory performance: A comparison of high-and low-aptitude children. Journal of Educational Psychology, 81, 306-312.

Spilich, G. J., Vesonder, G. T., Chiesi, H. L., & Voss, J. F. (1979). Text processing of domain-related information for individuals with high and low domain knowledge. Journal of Verbal learning and Verbal Behavior, 18, 275-290.

Stahl, A. S., & Jacobson, M. G. (1986). Vocabulary difficulty, prior knowledge, and text comprehension. Journal of Reading Behavior, 18, 309-324.

Stahl, A. S., Jacobson, M. G., Davis, C. E., & Davis, R. L. (1989). Prior knowledge and difficult vocabulary in the comprehension of unfamiliar text. Reading Research Quarterly, 24, 27-43.

Table 1 Ratio (and percentage) of studies that met criteria for internal validity (adapted from Lysynchuk et al., 1989).

Description of criterion	Studies that met criterion Ratio*	%
General design		
Control group present	16/16	100
Subjects were randomly assigned to conditions.	1/12	8
Subject mortality was approximately equal in treatment and control conditions.	16/16	100
Independent variables were explicitly described.	16/16	100
Dependent variables were explicitly described.	16/16	100
Dependent measures had face validity.	16/16	100
Hawthorne effects were unlikely.	16/16	100
Conclusions followed logically from the data.	16/16	100
Possible confounds		
Trained and control subjects exposed to same materials	11/11	100

Both trained and control subjects had equal time of exposure to materials.	10/16	63
Information was provided about time on task for both control and trained subjects.	9/16	56
The same experimenter provided treatment to all conditions	8/10	80
Absence of additional confounds	6/16	38
Measurement		
There were manipulation checks to determine that subjects did as instructed.	4/16	25
Alternate forms were used with repeated dependent measures.	1/1	100
There were no ceiling or floor effects.	6/6	100
Dependent measures were reliable.	2/5	40
Interrater reliabilities were reported.	5/13	38
Regression to the mean could be ruled out.	14/16	88
Statistics		
Probability of Type 1 error rate was controlled.	16/16	100

Unit of analysis was consistent with unit of treatment.	16/16	100
Correlation coefficients were computed within groups.	3/3	100
Data assumptions were discussed.	0/16	00
Cell size was reported.	8/16	50
Means were reported.	13/13	100
Standard deviations were reported.	7/13	54
Equal slopes treated in ANCOVA.	0/1	00
Information was provided as to type of ANOVA in unbalanced factorial designs.	0/4	00

*Ratio applies to studies for which the criterion was applicable. That is, the denominator of the ratio could be less than 16 when the criterion was not applicable, or when insufficient information was available to judge whether the criterion had been met.

Table 2 Ratio (and percentage) of studies that met criteria for external validity (adapted from Lysynchuk et al., 1989)

Description of criterion	Studies that met criterion Ratio*	%
Questions in the study were motivated by a theoretical or research base.	16/16	100
Characteristics of the sample were described.	3/16	19
Characteristics of the standardization sample for measures used in the study were similar to the sample participated in the study.	3/3	100
Information about the reading ability of subjects was given.	1/16	6
Information about the readability of text was given.	3/16	19
The study included a measure of delayed effects.	0/16	00

*Ratio applies to studies for which the criterion was applicable. That is, the denominator of the ratio could be less than 16 when the criterion was not applicable, or when insufficient information was available to judge whether the criterion had been met.